

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2002, 15:18:56 ; Search time 101.54 Seconds
(without alignments) 443.824 Million cell updates/sec

Title: US-09-052-089a-1
Perfect score: 2384
Sequence: 1 MPRALCTICSDFFDHSRDV.....VRVKTVPSELFOAKLDTFLMS 469

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	925	38.8	433	2 T30807	TRAF interacting p
2	204.5	8.6	506	2 F85016	probable RING zinc
3	187.5	7.9	263	1 S28261	centromere protein
4	183.5	7.7	1957	2 T38077	hypothetical coile
5	182.5	7.7	1290	2 A55094	chromosomal protei
6	182	7.6	425	2 T25457	hypothetical prote
7	179	7.5	1138	2 T24635	hypothetical prote
8	178.5	7.5	754	2 T51302	myosin heavy chain
9	178	7.5	477	2 JE0343	myosin heavy chain
10	177.5	7.4	1133	2 T22976	tear protein - rat
11	176.5	7.4	1156	2 E69444	hypothetical prote
12	176	7.4	1156	2 B70356	chromosome assembl
13	176	7.4	1509	1 A27232	myosin heavy chain
14	175	7.3	1690	2 T13030	microtubule bindin
15	172	7.2	1938	2 A59293	myosin heavy chain
16	172	7.2	2116	2 A26655	myosin heavy chain
17	171.5	7.2	1999	1 S21801	myosin heavy chain
18	170	7.1	532	2 A54871	M protein - strept
19	170	7.1	978	2 A70387	conserved hypotet
20	169.5	7.1	1744	2 JH0720	tanabin - African
21	169.5	7.1	1940	1 A24922	myosin heavy chain
22	169	7.1	472	2 S43554	plasmaogen-bindin
23	169	7.1	1935	1 S06006	myosin beta heavy
24	168.5	7.1	3187	2 JCS837	364K Golgi complex
25	168	7.0	959	2 A55913	transcytosis-assoc
26	168	7.0	1164	2 T24806	hypothetical prote
27	168	7.0	1325	2 T42722	male-enhanced anti
28	167.5	7.0	1132	2 T00259	hypothetical prote
29	167	7.0	853	2 T51505	hypothetical prote

30	167	7.0	879	2 A48575	paramyosin - nemat
31	167	7.0	1201	2 A35815	myosin heavy chain
32	167	7.0	1201	2 B35815	myosin heavy chain
33	167	7.0	2411	2 B32491	myosin heavy chain
34	166.5	7.0	555	2 C96667	unknown protein, 7
35	166.5	7.0	1931	2 A59234	slow myosin heavy
36	166	7.0	388	2 A49545	plasmaogen-bindin
37	166	7.0	484	2 S46489	M1 protein precurs
38	166	7.0	1738	2 T14867	interaplin - slime
39	166	7.0	1935	2 A59286	myosin heavy chain
40	166	7.0	2649	2 A40937	bulious pemphigoid
41	165.5	6.9	1295	2 T24587	hypothetical prote
42	165.5	6.9	1790	2 S67593	transport protein
43	165.5	6.9	1961	1 A61231	myosin heavy chain
44	165.5	6.9	2139	2 T18296	myosin heavy chain
45	165.5	6.9	2954	2 T14156	kinesin-related pr

ALIGNMENTS

RESULT	1	ALIGNMENTS
T30807	TRAF interacting protein - Fugu rubripes	
C:Species: Fugu rubripes		
C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000		
C:Accession: T30807		
R:Collage, A.J.; Clark, M.; Hawker, K.; Umrania, Y.; Wheller, D.; Bishop, M.; Elgar, F.E.S. Lett. 443, 370-374, 1999		
A:Title: Three receptor genes for plasmaogen related growth factors in the genome of		
A:Reference number: 220880; MID: 99148833		
A:Accession: T30807		
A>Status: preliminary; translated from GB/EMBL/DBJ		
A:Molecule type: DNA		
A:Residues: 1-433 <COT>		
A:Cross-references: EMBL:AJ010317; NID:e1355235; PID:e1355237; PIDN:CAA09084.1		
C:Genetics:		
A:Gene: TRIP		
A:Introns: 33/2; 52/3; 80/3; 94/1; 136/3; 168/2; 206/2; 235/3; 265/3; 310/1; 352/3; 3		
Query Match	38.8%;	Score 925; DB 2; Length 433;
Best Local Similarity	43.8%;	Pred. No. 4.5e-42;
Matches 197; Conservative 76; Mismatches 121; Indels 56; Gaps 5;		
QY	1 MPRALCTICSDFFDHSRDVAAIHGHTFHLOCLIOSFETAPSRTPCPOCRIOVGKRTIIN 60	
DB	1 MPRAOCTICSDFFDHSRDVAAIHGHTFHHECLVRFQAPRTKPCQCKROVSTRHIIS 60	
QY	61 KLFPDLAQEEENVLYDREFLKNELNDVNRALQSKDKERDSQVITIDTLDLEERNATVVS 120	
DB	61 RLTFVDGLDSSVGDPEPSLDNEIDRVAVNFSKREDRKQAKADDLMEIVELQRALEN 120	
QY	121 LQALGKRAEMICSTLTKOMRYLEEQODETRQAOEAGRLRSKMTQMEIILLQSOLEPV 180	
DB	121 LQKQVMEKEKLSALPTQMTYVLESQINDTFAAKEEVRRLRIKTKTFPSLDVYVLOGQAEV 180	
QY	181 EEMTRDMGVGQSAVEOLAVYCVSLKKEYENLKEARKASGEVADKLKRDLEFSSSKLOTVY 240	
DB	181 ESMITDMGIDQAAVEOLISYICISLKEYEDNLGGLKSSNMCEKLKREVLVTSNNKLKAL 240	
QY	241 SELDQALTELSAQKQDSADKREISSLKRLTLMQLEFLNLPVVAEFTVDRVLVE--SPAP 298	
DB	241 LELTKADDMKSLQNDLTNMEKEISSLKKEYEFLQELSLPTPTKRNEMAGLFFERCGAAP 300	
QY	299 ----VEVNEKLRRPSEFRDDIDLNATFVDVTPPARPSSOHGYEKLCLKLSHSPIDVP 353	
DB	301 PONSRAVSKASCLHPSPGNEIDIDLMYDVYTP-----DDVR 336	
QY	354 KTKCGRKRESQSLSGQSCAGEPEDELVGAFPLFVNATILGQKOPKRPSSSCSDVY 413	
DB	337 KRPKVSKRKMHL-----DSVILLRLRTKYKNLVSNQK-----FV 371	

QY 414 RTGFDGLGGRTRFIQPTDVTWIRPLPYKPK 443
 Db 372 QSGYDGLGGRTRKFIQPTRIHVRSEASKPE 401

RESULT 2
 F85016
 Probable RING zinc finger protein [Imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
 C:Accession: F85016
 R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
 Nature 402, 769-777, 1999
 A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A:Reference number: A85001; MUID:20083488
 A:Accession: F85016
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-506 <STO>
 A:Cross-references: GB:NC_001268; NID:g7267624; PIDN:CAB80936.1; GSPDB:GN00140
 C:Genetics:
 A:Gene: AT4G01270
 A:Map position: 4

Query Match 8.6%; Score 204.5; DB 2; Length 506;
 Best Local Similarity 21.5%; Pred. No. 0.00082;
 Matches 106; Conservative 88; Mismatches 189; Indels 111; Gaps 18;

QY 5 ALCTGSD---FPDHSRDVAHICHTFHLQCLIQSFETAPS---RTCPQCRIOVGKRT 57
 Db 10 ALCICYEDLKPVVENLOSISA--CGHVFEHLCLQOMFEYCPSYKRNCPICKQKSLKD 67
 QY 58 IINKLFDP-----LAQEEENVLD---ELFKENLDNVRAQLSQK----- 93
 Db 68 PC-RLYFOSSGNQTDSTASDKVGVIEDPVALRGEVRLBEGVONLTSALEAKKENVEV 126
 QY 94 -----DKERR-DSQVIIDTLRDLTEERNATVVSLOQALKAEMLCSTLKKOM 139
 Db 127 SDRKHQCNEDLKEDKVRWMALEIGISTTHLKLKSECCIQLNOCYKQLDERTVALANKEL 186
 QY 140 KYLEQODEFKQAOE--EAGRLNSKMKTMEOIEILLQS-----QLPEVEIMDMGVQ 191
 Db 187 ASKLVSDDLSEDDVRLKLLGNNAKTIDTLVRLSVLRNRSYELLAKCNOQLRGE 246
 QY 192 S-AVEQLAVYCVSLKKEYENLKEARKASGEVADLRKLDPSRSKLDQVYSELDOAKTEL 250
 Db 247 ARSEKELEKALEKEIKLKKRRELITTEERENALRDINYSK---KCSITEVSEPAIES 303
 QY 251 KSAQKDLQADKEIMSLKKRLTMLQETLNPVASETVDRVLVESPAPVEVNLKLRPSF 310
 Db 304 MSFRLMSDNK-----VEKISPPCKLEEKDGTFTIGSC-----LNGRDSF 346
 QY 311 RDDIDLNATDVTPPARPSSOHGYEKLCKLEKSHSPIQDVPAKTKGPKPK----- 363
 Db 347 VSRDSDVIDVDVDDVPTNTSGIRDMNTNIEKGDNSMVDIKINIKIDPSSVPSYNG 406
 QY 364 -----SOLSLGQSCAGEPPELVGAPPIVVRNAIIGQKPKRPRSSSSSKYVR 414
 Db 407 SGNTWSSSGTNRNLGRMSKHGERNE---ATP-----SLGGSYPRK-----DDLIS 448
 QY 415 TGFQDGLGGRTRFIQ 428
 Db 449 IGPDKGKGRIKIVLR 462

RESULT 3
 S28261
 centromere protein E - human
 N:Alternate names: centromere 312K protein; kinesin-related protein CENP-E
 C:Species: Homo sapiens (man)
 C:Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 19-Jan-2001
 C:Accession: S28261

R:Yen, T.J.; Li, G.; Schar, B.T.; Szilak, I.; Cleveland, D.W.
 Nature 359, 536-539, 1992
 A:Title: CENP-E is a putative kinetochore motor that accumulates just before mitosis.
 A:Reference number: S28261; MUID:93024922

A:Molecule type: mRNA
 A:Accession: S28261
 A:Residues: 1-2663 <YEN>
 A:Cross-references: EMBL:215005; NID:g29864; PIDN:CAA78727.1; PID:g29865
 C:Genetics:
 A:Gene: GDB:CENPE
 A:Cross-references: GDB:361164; OMIM:117143
 A:Map position: 4q24-4q25
 C:Superfamily: centromere protein E; kinesin motor domain homology
 C:Keywords: ATP; coiled coil; microtubule binding; mitosis; nucleotide binding; P-loop
 F:7-33/Domain: kinesin motor domain homology <KMO>
 F:86-93/Region: nucleotide-binding motif A (P-loop)
 F:486-2183/Domain: coiled coil #status predicted <COI>
 F:92/Binding site: ATP (Lys) #status predicted

Query Match 7.9%; Score 187.5; DB 1; Length 2663;
 Best Local Similarity 25.1%; Pred. No. 0.048;
 Matches 93; Conservative 66; Mismatches 133; Indels 79; Gaps 18;

QY 55 KRTIINKLFPDLAQEEENVLDREFLNKELDNVRAQLSQKDKER-RDSQVIIDTLRDLTEE 113
 Db 1632 KMTAVN-----ETQKMCGEIHLKEQFTQKLNLENITENIRLQI-----LHENLEE 1680
 QY 114 RNATVVSLOQALKAEMLCSTLKKOMKYLEQODET-----KQAOEENG-----RLRSKMT 165
 Db 1681 MR-SVTKERDRLRVE---ETLKVERDQLKENLRETTIRDLERQOEELKIVHMHKEHQET 1736
 QY 166 MEQIEELLQSLPPEVEIMDMGVGQSAV-----EQLAVYCVSLKKEYENLKEARKA 217
 Db 1737 IDKLRGIVSEKTNFISMQKDLNLSNDALKAQDLKIOEELRIAMHKLKEQDETIDKLRGI 1796
 QY 218 SGEVADRL---RDLDFSSRSKLDQVYSELDOAKLELSAOKDLQADK---EIMSLKKR- 270
 Db 1797 VSEKTDLSMKNQDLNLSNKKLQEKIQELKANHQTLTKLVNENPQKVSMEQQLKQI 1856
 QY 271 -----LTMQL-ETLNPVASETVDRL-----VLESPAVENVKLKRRSFDDI----- 314
 Db 1857 KDQSLTSLKLEIENLNAQELHLENLEMKSVMERDRLRVEETLKERQDLKESLQETK 1916
 QY 315 --DLNATFDVTPPARPSSOH-----GYEKLCK-----LEKSHSPIQDVPAK 355
 Db 1917 ARDLQIDQELKT--ARMLSKHEKETYDKLRKISERTIQISDIQKDLKSKDELQ---KK 1971
 QY 356 ICKGPRRESQL 366
 Db 1972 IQELQKKELQL 1982

RESULT 4
 T38077
 hypothetical coiled-coil protein - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T38077
 R:Connor, R.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
 submitted to the EMBL Data Library, April 1996
 A:Reference number: 221767
 A:Accession: T38077
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1957 <CON>
 A:Cross-references: EMBL:470690; PIDN:CAA94624.1; GSPDB:GN00066; SPDB:SPAC1F3.06c
 A:Experimental source: strain 972h-; cosmid c1F3
 C:Genetics:
 A:Gene: SPDB:SPAC1F3.06c
 A:Map position: 1

```

Query Match          7.7%; Score 183.5; DB 2; Length 1957;
Best Local Similarity 21.9%; Pred. No. 0.054;
Matches 76; Conservative 68; Mismatches 126; Indels 77; Gaps 11;

QY 70 EENVLDREFLNELNDVNRALQSKDEKRSQYIIDLRDLERNATVVSLOALQAE 129
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1386 EDNQLATNKLNDLHNLQDRIKLEDEVLKESKESLISLESLSNQPKESSLLDAKNELE 1445

QY 130 -MCSSTLKQKMKYLEQOQDETQAOEAGRLRSKMKTM-----EQIELL--LQSQLP 178
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1446 HMLDDTSRKNSLMKESTINSSLDKSFELASAVEKLALQKLHSESLSMENIKSOLQ 1505

QY 179 EVEEMIRDMGVGOSAAVEQLAVVCSLKEKEN-----LKEARK 216
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1506 EAKEKID---VDESTIOEDHEITASKNNYEGKNDKDSIIRDSLENIQDLNNLAEKS 1562

QY 217 ASGEVADKLKKDLFSSRSKL-----QTVYSELDQAKLELSAQKDLSADKEIMSLK 269
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1563 AVKRLSTEKESSEILOFNSRLADLEYHKSQVESELGRSKLKLAFTTEELQLAENERSLTT 1622

QY 270 KLTMLOETL---NLPPVASETVDRL-----VLESAPVPEVNKLK 306
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1623 RMIDLQNVQVDSIKSLSLSDTLRLSLSDVASLQKECKIKSNVYESLQDVLTSVQAR 1682

QY 307 RPSFRDDIDLNATFVDVTPPARSSSQH--GYEKLCKLEKSHSPID 351
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1683 NAELEDEVSRS---VDKIRRDRCHEHLSGKLKLL-----HSQLEE 1720

RESULT 5
A55094
chromosomal protein XCAP-C - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 02-Jun-2000
C:Accession: A55094
R:Hitano, T.; Mitchison, T.J.
Cell 79, 449-458, 1994
A:Title: A heterodimeric coiled-coil protein required for mitotic chromosome condensatio
A:Reference number: A55094; PMID:95042742
A:Accession: A55094
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1290 <HIR>
A:Cross-references: GB:U13673; NID:9563811; PIDN:AAA64679.1; PID:9563812
C:Superfamily: chromosome segregation protein SMC1
C:Keywords: chromosomal protein; DNA condensation; heterodimer

Query Match          7.7%; Score 182.5; DB 2; Length 1290;
Best Local Similarity 22.1%; Pred. No. 0.037;
Matches 100; Conservative 82; Mismatches 150; Indels 121; Gaps 20;

QY 15 DHSRDVA-AIHCGHTFHLOCLIOSFETAPSRTPQC-----RIOVGKRTII----- 59
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 620 DEKDVVAISSCGALDHI--VVDFTIDTA-----QECVNFLLKQNVGATFGLDKMKVWE 672

QY 60 ---NKL-----FPLDAQEEENVLDREF-----LKNELDNVRAQLSQKDEKRS 100
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 673 KGLNKTOTPEPIPLFPMVKVKKDQIKPAFTFALRDTIVANNLDQATRVAFQKDKRM-- 730

QY 101 QVIDTLRDLTLERNATVVSLOALGRAEMLSTLKR---QMKYLEQO----- 145
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 731 ---VVTILOGQIIEGSGMTGGCGKVMKGRMSSVMVELSDQDLQKMKKLTDTTRATEI 787

QY 146 QDETQAOEAGRLR-----SKMTMEQIELLQSQLPREVEMIRDMGVGQ 191
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 788 QDRRAHLEEEVAKLRQATREKMTFEKYTASLOSISEQVHLKAQVELEVNVAAPDK 847

QY 192 SAVROLAVVCSLKEKENLKE--ARKASGEV-----ADLKRDLSSRSKLTQTVSEL 243
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 848 NQCKMEKNETLTKKETKAEKAGKVAEAKRLHLKLIIVDNNHKLKRAQQDKLKVTKET 907

QY 244 DQ-----AKLELSAQKDLSA-----DKEIMSLKKLTMLOETL----- 278

```

```

      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 908 DECASTATKQVSKTRADRLKSEEAVALTEKEIYANAKSIELEDEKLKEKATYVA 967

QY 279 -----NLPPVASETVDRLVLESAPVPEVNKLKRPSFRDDIDLNATFVDVTPPARPS 330
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 968 NECKEACSLPEV--QDRHSRLQEIKAIOKEHALQREAL--NIRLNIE-QIDSHIAE-H 1022

QY 331 SSGHYEKLCKLEKSHSPIDVPEKKICKGPRKE 363
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1023 QSKIKYQKEITKISLHKIEDIPEEVLPGLAQE 1055

RESULT 6
T25457
hypothetical protein B0432.9 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Sep-2000
C:Accession: T25457
R:Henkhaus, J.; Wohldmann, P.
submitted to the EMBL Data Library, December 1996
A:Description: The sequence of C. elegans cosmid B0432.
A:Reference number: Z20038
A:Accession: T25457
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-425 <HEN>
A:Cross-references: EMBL:U08036; PIDN:ABJ7893.1; GSPDB:GN00020; CESP:B0432.9
A:Experimental source: strain Bristol N2; clone B0432
C:Genetics:
A:Gene: CESP:B0432.9
A:Map position: 2
A:Insertions: 64/3; 99/3; 165/1; 267/3; 350/2; 386/2
C:Superfamily: RING finger homology
F:188-238/Domain: RING finger homology <RRN>

Query Match          7.6%; Score 182; DB 2; Length 425;
Best Local Similarity 23.0%; Pred. No. 0.01;
Matches 65; Conservative 55; Mismatches 102; Indels 60; Gaps 10;

QY 3 IRALCTICSDFFDHSRDVA-AIHCGHTFHLOCLIOSFETAPSRTPQCRIQVGKRTIINKL 62
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 188 LQSGCSICFEDLQNDKISAIYCGHIYHNCISQMIAT--KRCPCSRRTVPKNGFEVKL 245

QY 63 FFDL-----AQEENVLDRFLKNELDNVNRALQSKDKRKRSQYIIDLRDLERN 116
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 246 FFDVQRMGGEAEKRPPEIDYRE-----EHYKLSLSTLVEQEK---LGTLTNTENKNLKD 294

QY 117 TVVSLQOAL-----GRAEMLSTLKQKMYLEQOODETKQAOEAGRLRSKMTMEQIELL 172
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 295 TVKSLKKIIRKNDKYQDELPKLQATINHLTISEEFAYLKRLEQESKNMLKTCFYLK 354

QY 173 LQSQLPREVEMIRDMGVGOSAVEQLAVYC-----VSIKKEYENLKEARKASGEVADKLK 227
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 355 TVH-----SSPADKOLGEYLKKNGLDTEKFEQQLAKSTYK--DLTDKRRE 397

QY 228 DLFSRSKLTQTVYSELDQAKLELSAQKDLSADKEIMSLK 269
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 398 -----AKEIQLKMEVOS---LKRRAQEDAAIKK 424

RESULT 7
T24635
hypothetical protein T07C4.10a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 01-Dec-2000
C:Accession: T24635
R:Buck, D.
submitted to the EMBL data Library, February 1995
A:Reference number: Z19915
A:Accession: T24635
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

```


RESULT 10

hypothetical protein F59A2.6 - Caenorhabditis elegans
T22976
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C/Accession: T22976; T23157
R/Lighting, J.
submitted to the EMBL Data Library, June 1994
A/Reference number: Z19645
A/Accession: T22976
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1133 <M1>
A/Residues: 1-1133 <M1>
A/Cross-references: EMBL:Z4801; PIDN:CAA84332.1; GSPDB:GN00021; CESP:F59A2.6
A/Experimental source: clone F59A2
R/Burton, J.
submitted to the EMBL Data Library, October 1995
A/Reference number: Z19700
A/Accession: T23157
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1133 <M1>
A/Cross-references: EMBL:Z6514; PIDN:CAA91344.1; GSPDB:GN00021; CESP:F59A2.6
A/Experimental source: clone K01A11
C/Genetics:
A/Gene: CESP:F59A2.6
A/Map position: 3
A/Intons: 13/2; 43/3; 107/3; 413/3; 492/3; 567/3; 635/3; 710/3; 738/3; 795/2; 1008/3; 1

Query Match 7.4%; Score 177.5; DB 2; Length 1133;
Best Local Similarity 20.5%; Pred. No. 0.059;
Matches 102; Conservative 103; Mismatches 169; Indels 123; Gaps 22;
QY 46 CPOCRIOVGR---RTIINKLFEDLAQEEENV-----LDREFLKNEIDNVRAQ 89
DB 57 CDALQAEVNAKALREIEQAKYDDVTKAEIRIGELSESKVLESEKQAFENEKEQREE 116
QY 90 LSKDKKRRS-QYIIDLTADLTLEFNATVVS-----LQALGKAMCSTLKKMKY 141
DB 117 QLAAMKRLSEONILDEVTKLEQSEEVLAANGAIOELTEKLESEKSTAKTELEA 176
QY 142 LEOQODETKAOEAGRLSKMKTM-----EOLELLIOQLPVEEMIRMGVGO- 191
DB 177 VSKKLDSSETSLKFEFSMIEAMKIQLINCEKQKDEAVELKQ- KLEVEKMSDVEVQKQ 235
QY 192 -----SAVEQLAVVCSLKEKEYENLEAKKASGEVADKLKDFSSRSKLTQTV- 240
DB 236 LLESTTSEMKQHAEAEIVKQ---LEFAQSS-----IENLKAENERN-LKTALSEDE 287
QY 241 -SELDQAKLELKSQKQLOASDKETMSLKKLTMLOET-----LNIPIV 283
DB 288 SSASISLTKQMEAKKLELSEKESSELRCQMRLQVHNAGODIQLKQTMLEMAKI 347
QY 284 ASETVDLTVESPAVENVNKLRRPSFRDIDLNATFDVDTPRPARPSSQHGVEKLCLE 343
DB 348 AKSTEDS-----KLAREQLAGELE-NAKEDL-----KVEE 377
QY 344 KSHSPIQ-----DVBR-KICKGRPKRESQSL-GGOSGAGEPDEELVGFPIFVNAI 393
DB 378 EKHGTGIRAGALDADAEKVKLEQLEQLESALESSQELASQKAKIQELELELQNA- 436
QY 394 LGQKQPRPSSSCSDVVRT---GPDGLGGRTKFIQPTDVTWIRPLPVKPKTK--VKQ 448
DB 437 -----QKRSEELLETANEMVRSLATLENSNTEILKQKLETLDEKLOAQROQTEKALTE 491
QY 449 RVKRVTVPSLFOAKLDT 465
DB 492 EINVLTT-SLAEKEQQT 507

RESULT 11

chromosome segregation protein (smc1) homolog - Archaeoglobus fulgidus
E69444
C/Species: Archaeoglobus fulgidus
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 02-Jun-2000
C/Accession: E69444
R/Klein, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod
R.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E
Glodek, A.; Zhou, L.; Overbeek, R.; Cocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A/Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaite, B.P.; Sykes,
Smith, H.O.; Woese, C.R.; Venter, J.C.
A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch
A/Reference number: A69250; MUID:98049343
A/Accession: E69444
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-1156 <K1>
A/Cross-references: GB:AE000995; GB:AE000782; NID:g2689318; PIDN:AAB89690.1; PID:g264
C/Superfamily: chromosome segregation protein SMC1

Query Match 7.4%; Score 176.5; DB 2; Length 1156;
Best Local Similarity 21.4%; Pred. No. 0.068;
Matches 74; Conservative 89; Mismatches 116; Indels 67; Gaps 11;
QY 55 KRTIINKLFEDLAQEEENVLD-----REFLKNED-----NVRALQSKDKREKD 99
DB 671 KERLSKIKYELQREKGLFAELNRAESLRQYDEVDRLTGMISELRNRISLDEKIRT 730
QY 100 SQVITDLRD-----TLERNATVVSLOALGKAB-----MLCSTLK 136
DB 731 ESGRIEELREISQSKREKENYISLKDYNKSLAMEBAIGELAEIETFRMRGSEVP 790
QY 137 KOMVLEQODETKAOEAGRLSKMKTMEOIFLLQSQLP-----VEEMIR 185
DB 791 KIVLELTKIKIEHQRNREILISIKTIESLEFRQELSSMQEQQVYLDEIKRIDELIR 850
QY 186 DMVGQSAVQGLAVVCSLKEKEYENLEAKKASGEVADKLKDFSSRSKLTQTVSELDQ 245
DB 851 TIEGRKARVEINSELELRKERELKELGKRDELDLQOLNMAEERKTEAETDR 910
QY 246 AKLELKAQKQLOASDKETMSLKKLTMLOETLNLPPVASETVDRVLESAPAVEVNLK 305
DB 911 LERIKIQKERTLEAESEIAETG-----VQPEMLPPL--EKVAK-VLD-----EVLVEL 958
QY 306 RRPFRDIDLNATFDVDTPRPARPSSQHGVEKLCLEKSHSPIQD 351
DB 959 ---STFGDVNLKAIQEEYEVKARDELVE---KMWLEKERADILD 998

RESULT 12

B70356
chromosome assembly protein homolog - Aquifex aeolicus
C/Species: Aquifex aeolicus
C/Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 02-Jun-2000
C/Accession: B70356
R/Dacker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
V.
Nature 392, 353-358, 1998
A/Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A/Reference number: A70300; MUID:98196666
A/Accession: B70356
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-1156 <AOF>
A/Cross-references: GB:AE000699; NID:g2983238; PIDN:AAC06839.1; PID:g2983243; GB:AE00
A/Experimental source: strain VF5
C/Genetics:
A/Gene: xcpC
C/Superfamily: chromosome segregation protein SMC1


```

Oy      358  KGPRESQJSLGGSCAEPDELVGAFPIFYRNALIGQK 398
          | : || | : : | | :
Db      1066 KAKTENLESTGTQTTIKDLQERLE-----ITNAELQHE 1100

```

RESULT 15
A:59293
skeletal myosin heavy chain - domestic rabbit
C:Species: *Oryctolagus cuniculus*
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 08-Sep-2000
C:Accession: A59293
R:Meda, K.; Hostilnova, E.; Roessc.Kleinkauf, A.; Schuster, H.; Gasperik, J.; Wittinghofe
submitted to GenBank, July 1995
A:Description: Isolation, sequencing of myosin heavy chain cDNA from rabbit skeletal mus
A:Reference number: A59293
A:Accession: A59293
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1938 <MAE>
A:Cross-references: GB:032574; NID:g940232; PIDN:AAA74199.1; PID:g940233
A:Experimental source: strain New Zealand White; cell type skeletal muscle fiber type I
C:Genetics:
A:Gene: MHC
A:Superfamily: myosin heavy chain; myosin motor domain homology
F:88-769/Domain: myosin motor domain homology <MMO>

RESULT 16
A26655
myosin heavy chain [similarity] - slime mold (Dictyostelium discoideum)
N:Contains: myosin ATPase (EC 3.6.1.32)
C:Species: Dictyostelium discoideum
C:Date: 05-Oct-1988 #sequence-revision 05-Oct-1988 #text-change 02-Feb-2001
C:Accession: A26655, A24728, S00250
R:Warrick, H.M.; De Lozanne, A.; Weinand, L.A.; Spudich, J.A.
Proc. Natl. Acad. Sci. U.S.A. 83, 9433-9437, 1986
A:Title: Conserved protein domains in a myosin heavy chain gene from Dictyostelium discoideum
A:Reference number: A26655; MUID:87092266
A:Accession: A26655
A:Molecule type: DNA

A:Residues: 1..1116 <W>
A:Cross-references: GB:MI14628; GB:MI1938; NID:9167834; PIDD:AAA33227.1; PID:9167835
R:Delozanne, A.; Lewis, M.; Spudis, J.A.; Leitwand, L.A.
Proc. Natl. Acad. Sci. U.S.A. 82, 6807-6810, 1985
A:Reference number: A24728; MUID:86016788
A:Accession: A24728
A:Molecule type: mRNA
A:Residues: 2035-2116
R:Wagler, G.; Noegel, A.; Scheel, J.; Gerisch, G.
FEBS Lett. 227, 71-75, 1988
A>Title: Phosphorylation of threonine residues on cloned fragments of the Dictyostelium
A:Reference number: S00250; MUID:86112226
A:Accession: S00250
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1734-1893 <W>
C:Comment: The rod domain is highly periodic, containing a pattern of 7-residue repeats
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; ATP; coiled coil; hydrolase; nucleotide binding; P-loop; P
F:1-818/Domain: globular head <HD>
F:89-747/Domain: myosin motor domain homology <MOT>
F:179-186/Region: nucleotide-binding motif A (P-loop)
F:819-1116/Domain: alpha-helical rod <ROD>

```

RESULT 17
S21801
myosin heavy chain, neuronal [similarity] - rat
N|Alternate names: myosin II
N|Contains: myosin ATPase (EC 3.6.1.32)
C|Species: Rattus norvegicus (Norway rat)
C|Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 19-Jan-2001
C|Accession: S21801; P00013; S18134
R|Sun, W.; Chantler, P.D.
J. Mol. Biol. 224, 1185-1193, 1992
A|Title: Cloning of the cDNA encoding a neuronal myosin heavy chain from mammalian brain
A|Reference number: S21801; MUID:92235856
A|Accession: S21801
A|Molecule type: mRNA
A|Residues: 1-1999 <SUN>
A|Cross-references: EMBL:X62659
R|Sun, W.; Chantler, P.D.
Biochem. Biophys. Res. Commun. 175, 244-249, 1991
A|Title: A unique cellular myosin II exhibiting differential expression in the cerebr

```

A:Reference number: PN0013; MUID:91151356
 A:Accession: PN0013
 A:Molecule type: mRNA
 A:Residues: 1914-1998, 'T' <SU2>
 A:Experimental source: Brain
 C:Superfamily: myosin heavy chain; myosin motor domain homology
 C:Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; nucleotide
 F:84-763/Domain: myosin motor domain homology <MOT>
 F:174-181/Region: nucleotide-binding motif A (P-loop)
 F:541-575/Region: actin binding #status predicted
 F:653-675/Region: actin binding #status predicted
 F:836-1276/Domain: coiled coil #status predicted <COI>
 F:1277-1999/Region: light meromyosin
 F:125/Modified site: N6,N6,N6-trimethyllysine (lys) #status predicted
 F:180/Binding site: ATP (lys) #status predicted
 F:693,703/Active site: Cys #status predicted
 F:1916/Binding site: phosphate (Ser) (covalent) #status predicted
 F:1943/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 7.2% Score 171.5; DB 1; Length 1999;
 Best Local Similarity 18.7%; Pred. No. 0.24;

Matches 83; Conservative 65; Mismatches 140; Indels 135; Gaps 14;

QY 48 QCRIOVGKRTINKLFFDLAEOEENVLDREF-----LKNELDNVRAQLSOK 93
 Db 1072 ELKMQGLAKKE--ELQNALRAVEEELAAQKNMALKKIRELSQISELDDELSEFASRNKA 1129
 QY 94 DKEKRSQVVIIDTLRDLTEERNATVSLQALGAEMLCSTLKKOM----- 139
 Db 1130 EKOKRDGEELEALKTELELDLTSTAAQELRSKREOVNLLKKTLEBEAKTHEAQIQEM 1189
 QY 140 -----KYLEQOODETKQAQEEAGRLRSKMTME-----QIEILLQS-----QLPEV 180
 Db 1190 ROKHSQAVEELAEOLDEQTKRVANLEKAKOTLENERGELANEVNVLLQGGDSHKKRKV 1249
 QY 181 EEMIRDGCV---GQSAVEDLAAYCVSLKKEYENLKEARKASGEVADKLRRKDFSSRSKL 236
 Db 1250 EAQOQELQVFNENGERVTELDADKVTALQVLELDVNTGLLSQSDSKSLTKRDFALSQSL 1309
 QY 237 QTVYSELDAQKLELSAQKDLQSDAKIMSLKKKLTMLQETLNPVAVSETVDRLVLESP 296
 Db 1310 Q-----DTQELQDENRQKLSLTKLQVDEBKN----- 1338
 QY 297 APVEVNLKLRPSFRDDIDLNATGFVDTPPARPSSOHGYEKLCLERS---HSPIDQV 352
 Db 1339 -----SFRQLE-----EEEBEAKNN-----LEKQIATLHAQVADM 1369
 QY 353 PKKI-----CKGRKESQSLSGQSCAGEPDEELVGAFFIVNAILGQKQ----- 398
 Db 1370 KKKMEDVSGLETAEEVKRKLQKDLBGLS---QRHEEKVAAYDKLEKTKRRLQOELDDL 1426
 QY 399 -PKRPRESSCSKDVATGFDGL 420
 Db 1427 VDLDHQOSACNLEKKQKFDOL 1449

RESULT 18

S54871
 M protein - Streptococcus sp.
 C:Species: Streptococcus sp.
 C:date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 26-Aug-1999
 C:Accession: S54871
 R:Podbielski, A.; Melzer, B.
 submitted to the EMBL Data Library, June 1991
 A:Reference number: S54871
 A:Accession: S54871
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-537 <POD>
 A:Cross-references: EMBL:X60097; NID:g840905; PIDN:CAA42693.1; PID:g840906
 C:Superfamily: M5 protein

Query Match 7.1% Score 170; DB 2; Length 532;
 Best Local Similarity 21.5%; Pred. No. 0.059;
 Matches 82; Conservative 67; Mismatches 141; Indels 92; Gaps 12;

QY 55 KRTIIN-----KLFFDLAEOE-----EENVLDREFLK-----NELDNVRAQ 89
 Db 159 KKNVNDNSYNTKREVDLIEELGKKLKENODLEBKDKDEFFLGLTRYINLDDLKLG 218
 QY 90 LS-----QKDKERDSQVVIIDTLRDLTEERNATVSLQALGAEMLCSTLKK 137
 Db 219 LINDINIDLKHELOEKQKAEADROT--LEAKKALEEKQISDASRSLSLRDLDASBEAK 277
 QY 138 QMYLYEEOOQETQAOQOEAGR--LRSKKWTMEQIEILLQSLQVPEE-----MIR 185
 Db 278 QLENEYOKLEEEKQISDASQSLRDLDASREKKQLEAYOKLEBQNTISEASRGLR 337
 QY 186 DMGVGQSAVBQLAAYCVSLKKEYENLKEARKASGEVADKLRRKDFSSRSKLQTVSELDQ 245
 Db 338 DLDSAREAKQKQVEKDLANLAEALDKYEEKQISDASRKLGRDLASR-----E 386
 QY 246 AKLELSAQKDLQSDAKELMSLKKLTMLQETLNPVAVSETVDRLVLESPAPVEVNLK 305
 Db 387 AK---KQVEKALEEANSKLALEKLNKELEES-----KLTLEKKAELQAKLEA 432
 QY 306 RRSFRRDDIDLN-----TFVDTPPARPSSOHGYEKLCLERSHSPIDQVPPK 355
 Db 433 EAKALEKLEKQAELEKLAKRAGKASDQTPAEATPGN-----KVYPGQAPQACTKPN 485
 QY 356 ICKGRKESQSLSGQSCAGEP 377
 Db 486 QNKPKMETKROLPSTGEATNP 507

RESULT 19

A70387
 conserved hypothetical protein aq_1006 - Aquifex aeolicus
 C:Species: Aquifex aeolicus
 C:date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 02-Feb-2001
 C:Accession: A70387
 R:Becker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
 Nature 392, 353-358, 1998
 A:title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
 A:Reference number: A70300; MUID:98196666
 A:Accession: A70387
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-978 <AOF>
 A:Cross-references: GB:AE000718; NID:g2983504; PIDN:AA07092.1; PID:g2983515; GB:AE00
 A:Experimental source: strain VF5
 C:Genetics: aq_1006
 A:Gene: aq_1006
 C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032

Query Match 7.1% Score 170; DB 2; Length 978;

Best Local Similarity 22.6%; Pred. No. 0.12; Mismatches 125; Indels 76; Gaps 13;

QY 41 APSRTCPQCRIOVGKRTIIN-----KLFFDLAEOEENVLDREF-----LKN 82
 Db 481 SPDDTCTPCVGGIIRGKALEVNDADGISELKNAKLEKEKEREIDTLTKLYAQKINSLEKE 540
 QY 83 LDNVRAQLSOKDKERDSQVVIIDTLRDLTEERNATVSLQALGAEMLCSTLKKOMKYL 142
 Db 541 MEKLRNEVEELRKE-----IPENLKER---IKLELELIEREKLEHKLNKYRKAL 587
 QY 143 EEOODETKQAQEEAGRLRSKMTMEQIEILLQSL--PEVEEMIRDMGVGQSAVQSAVEQLAYVC 201
 Db 588 EDRQKQKEEAQAKLRAQTELELLKE-KIREKSRLVKEKRELYR-----VERLEDYE 638

A; Cross-references: EMBL:Z32678; NID:g474769; PIDN:CA83589.1; PID:g1333838
C; Superfamily: M5 protein

Query Match	7.18;	Score 169;	DB 2;	Length 472;
Best Local Similarity	21.98;	Pred. No. 0.058;		
Matches	82;	Conservative	64;	Mismatches 132;
			Indels	96;
			Gaps	12

QY	65	DLA0E-EENVLDREFLK-----NELDNPAAOLG-----QKKEKERSOVII	104
Db	118	DLNODLEKLDKREFYGLGETRTIINELDLKIGQNTINIDLKHELEBOKKAENDROT-L	176
QY	105	DTLRDLEENATVATVSLQOALGKAEMCLSTLKQMKYLEOQODETKQAQEAAGR-LRSM	163
Db	177	EAEKALEEERQISDASRQSLRDLDSAREAKQYLEAEYQKLEERQISDASRQSLRDL	236
QY	164	KTMEQIELLOSQLEPVEE-----MIRMGVQSGSAVELOAVYCVLSLKEEYENLK	212
Db	237	DASRAKKQLEAEYQKLEERQISEASRKGLRDLDSAREAKQ-----LEAEHOKLE	289
QY	213	EARRASGEVADKLKKRDLEFSR-----SKQIYVYSELDAKLE-----	249
Db	290	EONKTISEASRKGLRDLDSAREAKQYOEKDLANLTAEIDLVKKEEQISDTSRKGLRDL	349
QY	250	-----LKSQOKDLOSADKEIMSLKKLTJMOETLNPVASETVDRLVLESPAPVEYNLK	304
Db	350	ASREAKKQYEAQALEANSKTLAALDKLKELEES-----KLTLEKKEKLEQAKLE	398
QY	305	LRRPSFRDDIDLNA-----TFDVPDPPAPRPSSGHGYEKKCLEKSHSPIDVPR	354
Db	399	AEAKALKETIKLAQAEELAKTLRAGTASDSQPRPEATPGN-----KVYRGQAPQAQGTKP	451
QY	355	KICGPKRESQSL	368
Db	452	NQNEPKMETKROL	465

RESULT 23

S06006

myosin beta heavy chain, cardiac muscle [similarity] - rat

N:Contains: myosin ATPase (EC 3.6.1.32)

C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 19-Jan-2001

A:Accession: S06006; S07536; 167441; A02989

R:Krafft, R.; Bravo-Zehnder, M.; Taylor, D.A.; Leinwand, L.A.

A:Title: Complete nucleotide sequence of full length cDNA for rat beta cardiac myosin heavy chain

A:Reference number: S06006; MUID:90016823

A:Accession: S06006

A:Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-1935 <RRA>

A:Cross-references: EMBL:X15939; NID:g56656; PIDD:CA934065.1; PID:g56657

R:McNally, E.M.; Krafft, R.; Bravo-Zehnder, M.; Taylor, D.A.; Leinwand, L.A.

J. Mol. Biol. 210, 665-671, 1989

A:Title: Full-length rat alpha and beta cardiac myosin heavy chain sequences. Comparison of the alpha and beta chains

A:Reference number: S07535; MUID:90133919

A:Accession: S07536

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-950, 'RK', 953-1935 <MCN>

R:Mahdavi, V.; Lompre, A.M.; Chambers, A.P.; Nadal-Ginard, B.

Eur. Heart J. 5, 181-191, 1984

A:Title: Cardiac myosin heavy chain isozyemic transitions during development and under pathological conditions

A:Reference number: I53305; MUID:85179510

A:Accession: 167441

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1871-1935 <RBS>

A:Cross-references: GB:W32698; NID:g205598; PIDD:AAA41659.1; PID:g205559

R:Mahdavi, V.; Petrasamy, M.; Nadal-Ginard, B.

Nature 297, 659-664, 1982

A:Title: Molecular characterization of two myosin heavy chain genes expressed in the adult rat heart

A; Reference number: A02988; MUID:82220036

A;Accession: A02989
A:Molecule type: MB

A; Residues: 1524-15

A; Cross-references: GB:J00752; NID:g205577; PIDN:AAA41654.1; PID:g205578

C;keywords: actin binding; ATP; cardiac muscle; coiled coil; heart; hydrophobicity; myosin heavy chain; myosin motor domain homology; superfamily; myosin heavy chain; myosin motor domain homology

F:88-766/Domain: myosin motor domain homology <MMOT>

E:548-585/Region: actin binding #status predicted
F:178-185/Region: nucleotide-binding motif A (P-loop

F:655-677/Region:	actln binding	#status	predicted
1,370-200/Region:	actln binding	#status	predicted

F;839-1935/Domain: coiled coil #status predicted <CO
E:839-1279/Region: S2

F;1280-1935/Region: light meromyosin

F:129/Modified site: N6,N6,N6-trimethyllysine (Lys)

F:184/Binding site: AMP (Lys) #status predicted
E:695.705/Active site: Cys #status predicted

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

Querv Match 7 19. Score 169. DR 1:

Query match	7.1%;	Score 105;	DB 1;
Best Local Similarity	22.5%;	Pred. No. 0.32;	

Matches 65; Conservative 62; Mismatches 106

0Y 31 LOCLIOSFETAPSRTPCPOCRIOVGKRTIINKLTFDLAOEEN

Query Match	7.1%;	Score 169;	DB 1;	Length 1935;
Best Local Similarity	22.5%;	Pred. No. 0.32;		
Matches 65;	Conservative 62;	Mismatches 106;	Indels 56;	Gaps 9;

Oy	31	LOLOSPEPAPSTCQCHIOYGRITIIINKLPEFDLOJEEBVLUREF-----	78
Db	1143	LEELSERLEEEGAGTSTVQIEMNNKKREEREFQKMRDL-----EELATLOHEATAALRRKHADS	1199
Oy	79	--LKNELDNVRAOLSOKEDEKRDQVITIDTLRDLERNATVVSLOALGAEMICSTL	135
Db	1200	VAEIGEOIDNLOPVKOKLEKSEFPLEDDVTSNMEO-----IIRKANLEKICRTL	1252
Oy	136	KKOKNYLEQOODDEFKQOEEAGRLRSMTK-----MEQIELLOSLP-----	178
Db	1253	EDQNNHEHRSKLEEFORSVNDLTQORAKLOTEGSELSRQLEDEBAL-SOUTRGKLTYYQO	1311
Oy	179	-----EVEEMIRDMGOSAEQVLAVVYCSLKEEYNLEKARKASGEVADLRKDLFSS	232
Db	1312	LEDLRQLEEEVKKNNALAHALQSARHDDCLAREQEETFEAKAELOVRLSKANSEAVQW	1371
Oy	233	RSKSLQI-----VSELDQAKLELKSQAKDOSADKELMSLKKULTMIOER	277
Db	1372	RKKEETAIORTLEFEAKKRL--AOR-LQDAEEAAVEANACSSISLEK	1417

RESULT 24

JC5837

364k Golgi complex-associated protein - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 20-Jun-2000

C:Accession: JC5837

R:TokI, C.; Fujiwara, T.; Schda, M.; Hong, H.S.; Mismu, Y.; Ikehara, Y.

C:Title: Identification and characterization of rat 364-kDa Golgi-associated protein

A:Reference number: JC5837, MUID:98093490

A:Accession: JC5837

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-3187 <TOK>

A:Cross-references: DDBJ:D25543; NID:9516825; PIDD:BA05026.1; PUD:9516826

C:Comment: This protein plays a role in the formation and maintenance of the character

C:Superfamily: giantin

F:49-549,624-1176,1238-1707,1763-3114/Domain: coiled-coil leucine zipper #status pred

F:3165-3187/Domain: membrane anchor #status predicted <MAD>

Query Match	7.1%;	Score 168.5;	DB 2;	Length 3187;
Best Local Similarity	19.3%;	Pred. No. 0.61;		
Matches	89;	Conservative	84;	Mismatches 160;
				Indels 127;
				Gaps 13;

```

QY 65 DLQEEENVLDREFLKN--ELDNRALQSOKDKER-----DSQYIDTLRDLFEER 114
      :| : : | | | | | : : : : : : : : : : : : : : : : : : : : : :
Db 1310 ELSSSQLKADLEHLKTLQLELEFLQKHVQSGKEEFSYLVGOLGKEKQTLTVQTEMEQ 1369

```

QY	115	NATVNSL-----	COALGKAMLCSTJCKOMKYUEOQ	146
		: :	:	
Db	1370	ERLIKALHTLOEMQAKENHERLKQVOYETICELKQKOLEEBSKAKQOLQRRKLQALISR	142929	
QY	147	----DETQAOEBAGRLRSM-----	KTMOEILLQSOQPEVEEMIRMGVQASAVEOLA	198
		: :	: : : :	
Db	1430	KEALKEKNSLQEOBSSARDAVEHLTRSLADVSQYSVQNOEKDALLGKALLQOEOROLI	14899	
QY	199	V-----YCVSLK-----		205
Db	1490	VEMDKSLLENQSLGSGCESLKLALGLTEDEKLEKMLESVRCSKIATSEWQEKHELO	1549	
QY	206	KEYENLKEARKASGEVADKRNDLPSSRSKLOTVYSELDOAKLEKSNOKDQASDKRIM	265	
		:	: : : :	
Db	1550	KEYEVLLOSTIENNSNERQIHQWESVROKQOEYAKLESASDSKRRERKQLOQDEOME	16099	
QY	266	SLKKKL-----TMOETLNLPPVASETVORLLESFAPVEV-----	EKLRRPSFRDID	315
		: :	:	
Db	1610	EMKEKMKKFAKSNQOKITLEL-----	BEENDRLAEHQPVQGANESMEDALLSSNASLKELE	16655
QY	316	LNATFVDVDPARPASSQHGYEYKCLCEKXSHPIQDVPKKICGFRKESQSLSGQSCAG	375	
		: :	:	
Db	1666	RITLEKYLTSKE-----	FEALMAEKNTLSSEETRNKLQLOVEAOELKOASLFTTEKSD	1716
QY	376	EPDEELVGAFPIFVRNALIGOKOPKRRPSSSCSKVAPT	415	
		: :	: : : :	
Db	1717	EPKDVIE-----VTEAVVKGSOBQJSLSENKLDEAEAT	1751	

RESULT 25

A55913

transcytosis-associated protein p115 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 05-Nov-1999

C:Accession: A55913

R:Barroso, M.; Nelson, D.S.; Sztul, E.
Proc. Natl. Acad. Sci. U.S.A. 92, 527-531, 1995

A:Title: Transcytosis-associated protein (Ttrp)/p115 is a general fusion factor required

A:Reference number: A55913; MUID:95132633

A:Accession: A55913

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-959 <BAR>

A:Cross-references: GB:U05589; NID:g558474; PIDN:AAC52151.1; PID:g558475

C:Keywords: membrane fusion; membrane trafficking

Query Match	7.0%	Score 168;	DB 2;	Length 959;
Best Local Similarity	20.2%	Pred. No. 0.15;		
Matches	76;	Conservative	56;	Mismatches 93;
				Indels 152;
				Gaps 11;

```
QY      52 QVGRTTINKLEF-----DLAGEENVLDREFLK--NEDNY--RRLQS 91
       :::| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      571 RIGENIENTEKGFISKHELLPRASQKOPNPPEFYIIFDHEFLKLVEGYITTAIK 630
       :::| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      92 QDKREKRDSOYLIDLTDLTEERNATVSLQALGKAEMCLSTLKKOMKYLEQODE-- 148
       ::::: | | | | | | | | | | | | | | | | | | | | | | | | |
Db      631 SEEDKKEEV----KTLEQHNDINITYHKNNVIREDQLDEELKOQVSTLKCQNEQLQT 685
       :::| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      149 --TRQAQ-----EENGRLSRKMKTYMEOI 169
       :::| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      686 AVTQASAIQIOAHKDQYNILKVLQKDNHOGSHSDGAQVNGIQPBEISRLEETIELRSH 745
       :::| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      170 ELLQOSOLPEVEENIRDM-----GVGOSAV-----BOLAYV----- 200
       :::| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      746 QVLLQSLAEKDVIYENLRSSQVSGMSEBALATCSPDAEQVALEKOBLSALKSQLSQS 805
       :::| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      201 -----CVSLKREYENLKEARK 216
       :::| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      806 LEITRLQENRELQORAFETLAKSVPEGESELYPAKTTVDGRHSALLDETLELNKEIK 865
       :::| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      217 ASGEVADKLKRDLFSSRSKLOTIVSELDQAKLELKSQKD-----LQSADEIMSLK 268
       :::| | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Db 866 ALSEERSIQQLDSSNSTAIILQETENDKLYLEVTDSEKQDDLLVLLADQDKILSLKS 923

Qy 270 KLTMLOETLNLPPVASE 286

Db 926 KLKDLGH-----PVEEE 937

RESULT 26
T24806
hypothetical protein T10G3.5 - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T24806
R/Burton, J.
submitted to the EMBL Data Library, October 1996
A/Reference number: Z19937
A/Accession: T24806
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1164 <MW>
A/Cross-references: EMBL:Z81118; PIDN:CAB03330.1; GSPDB:GN00023; CESP:T10G3.5
A/Experimental source: clone T10G3
C/Genetics:
A/Gene: CESP:T10G3.5
A/Map position: 5
A/Introns: 15/3; 73/3; 387/3; 412/3; 708/3; 818/3; 894/3; 982/3; 1080/2

Query Match	7.0%;	Score 168;	DB 2;	Length 1164;
Best Local Similarity	24.3%;	Pred. No. 0.19;	Indels 18;	Gaps 8
Matches 62;	Conservative 66;	Mismatches 109;		

```

QY      31  LQCIJSEETAPNSTCPOGRCIOLGKFTIINKLFPDLAQEEENVLDREFLK--NELDNVRAQ  89
      667  LEAVSSKDTIVTSKSEBSLSLKOKLTENSFPIELVAVQEKYSNLSSEKQOEVENLME  726
Db
QY      90  LSKRD--KERKD--SOVITIDTLRDTLLEERNATVSVLQALCAKEMLCSTLKKOKYLE  143
      727  MRDKEAHHKTKRDFEFQOMLRN--QEDNEBASSTLKSVQOLMKKEKETSSEKKNOLISVY  784
Db
QY      144 QOOQETQAOQENARLASKKMTQOEIILLQOSLPVEEMIRMGVQASAVEDOLAVYCS  203
      785  SQDEELKTEVERL--INSEKTOIEIKLASAYATTQO--ROBLATFS--ESLRTJECEN  838
Db
QY      204 LKKEEYENLKEARKASGEVADKLKRDLPSSRSKQTVYSELDOAKLELKSQKODASDKR  263
      839  LNSKIQSIIEESRRHAEKGSNLEMRITERSRLE--KDIIEERESTIQSIQEALETRKDN  895
Db
QY      264  TMSLKKKLTMTQETLL  278
      896  IESLKTQTORVEDEL  910
Db

```

RESULT 27
T42722
male-enhanced antigen-2 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000
C:Accession: T42722
R:Kondo, M.; Sutoh, S.
DNA Seq. 7, 71-82, 1997
A:Title: Cloning and molecular characterization of cDNA encoding a mouse male-enhanced
A:Reference numbers: Z22242; MUID:97127683
A:Accession: T42722
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1325 <R0N>
A:Cross-references: EMBL:J78270; NID:d1096175; PID:d1020389; PIDN:BAJ9612.1
A:Experimental source: strain CD-1
C:Function:
A:Description: supposed to play some role for spermatogenesis
;Keywords: leucine zipper

Query Match	7.0%;	Score 168;	DB 2;	Length 1325;
Best Local Similarity	20.6%;	Pred. No. 0.23;		
Matches	77;	Conservative 65;	Mismatches 130;	Indels 102;
			Gaps	13;

```

0Y      68  OEEENVLDREFLKNEID-----NVRQLOSKDKKRSQVYIIDTLRDLRENNRTVYSL 121
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      966  REHNSILETALAKREADVQLNLYQVAVYLOKRKEEDROMKOLVQALOVSLEKEKMEVNL 1025
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      122  OQALCKA-----EMLCSTLKKOMKYLEQO----QDET 149
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1026  KEQMAARIIEAGHNRRHFKAATLELSYVKKELQAKENHVLQTLQAEVDELQIDQDKHSQDI 1085
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      150  KOAOEAGRLSKSKMTMEQIEILLQSOQLPEVEEMTRDMGVGQSAVEQLAVYCVSLKKEYE 209
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1086  AOFQTELEAFATOLO-----LLOKKTL--DEOMSQOPTGQEMEDTLKWELODKERETIO 1135
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      210  NLKEKRAKASGEVADLRKDLFSSBSKLOTVYSELQOAKLELKSACKDLOSACKELMSLKK 269
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1136  SILQOOLDLITEQG--KKELBGTQOTLOTIKSELEWQVQDLSJETQDKFMLOAKSELN 1192
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      270  KL-TMLOET-----LMLP--PVASETVRLVLESFA 297
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1193  NMKTLTLQNOQOLKDLRRGAAKKREKGBSSNSSPATPIKIDCCVPASLLLELLRPPA 1252
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      298  ---PVEYNLKLRRSPFRDDID-LNATFDVDTPPARPSSSOHQYIEKLCLEKXSHPIDOV 352
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1253  VSKPEPLK-NLNCLLOQLKQEMDSLOROMEHITIVHESLSSWAQYEAAPAHAN----- 1305
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      353  PKKICKGPRKESQL 366
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1306  -----PRGDTKL 1312
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 28
T00259
hypothetical protein KIAA0477 - human
C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: T00259
R:Sei1, N.; Ohira, M.; Nagase, T.; Ishikawa, K.; Miyajima, N.; Nakajima, D.; Nomura, N.
DNA Res. 4, 345-349, 1997
A:Title: Characterization of cDNA clones in size-fractionated cDNA libraries from human
A:Reference number: 214085; MUID:98116662
A:Accession: T00259
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1132 <SEK>
A:Cross-references: EMBL:AB007946; NID:g3413915; PIDN:BAA32322.1; PID:g3413916
A:Experimental source: brain
C:Genetics:
A:Note: KIAA0477
A:Keywords: alternative splicing

```

Query Match 7.0%; Score 167 5; DB 2; Length 1132;
Best Local Similarity 22.8%; Pred. No. 0.2; Mismatches 139; Indels 87; Gaps 21.
Matches 91; Conservative
QY 24 HCGHTE-HLQCLIOSFEFTAPSRFCPOCHRYQVKRTIINKLFDFDLQAEENVIDREF-LKN 81
   - - - - - : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 601 HLNLSLSKELLQEEF-----RELLQYRDNSSDKTLEAENML-LKLRORIHDKAAALER 653
   - - - - - : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 82 ELDNVRAQLSQDKDEKRDQVIL-----DTRDRLTEERNATVVSLOALG-----KAEM 130
   - - - - - : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 654 AIDEKFSALEREKELRLQRLAVRERDHDLERLDVLSNENATWQSMSSILAKGLEVEQ 713
   - - - - - : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 131 LCLSTLKKRMKTLFQOQODETK-----QAQEEAGLRKSKMTMBEIIELLOSQLEVEE----- 182
   - - - - - : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 714 L-STTCOMQWLKEM-ETKFSRWKQEDQS-----TIQOLQSLSDRNKEVEDLSNT 763
   - - - - - : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 183 MIRDMGVQSAV-EDLAVVYVSLKKEYNELKERRASGEVADKLKRDLEFSK----- 235
   - - - - - : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

[illegible]

RESULT 29
T51505
hypothetical protein F5E19_70 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000
C:Accession: T51505
R:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.;
submitted to the Protein Sequence Database, August 2000
A:Reference number: 225394
A:Accession: T51505
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-853 <SAT>
A:Cross-references: EMBL:AL391147
A:Experimental source: cultivar Columbia; BAC clone F5E19
C:Genetics:
A:Map position: 5
A:Introns: 6/2; 79/3
A:Note: F5E19_70

Query Match	7.0%	Score 167	DB 2	Length 853
Best Local Similarity	21.9%	Freq. No. 0.15		
Matches	93	Conservative	80	Mismatches 141; Indels 110; Gaps 19
QY	LAOEEENVLDREFLNENLDNVRAOLSOXDKERDRSOVIITDTRLRLEENATVVSLOAL	125		
DB	382 VAKQKE---DLEVSFORLGSVEEYSKNEKEVEKLKSELETYK---EKNRLAKKEQDAT	435		
QY	126 GKAEMLCSTLKKOMKYLE---OOODETRQAOB-----EAGRLLSKMKT-----	165		
DB	436 SRVQRLSBEKSKMLSDLESSEKBEESKKAEMSLASALHVESSSECRRELKILLSGDHXY	495		
QY	166 ---MBQITLLILOSULPEVBEIM---RDMGCGQSAVEDQALVCVSLKETEENLKEA----	214		
DB	496 ETQIDDLRLVITKATREKYEENMLDEARHEIDVLVSAAVEQTKKHFESSKDWME--MKENALVN	554		
QY	215 --RKASGVAAKRLKRDLLSSRSKLTQTVVSELD-----QAHLTELKSA-----OKDIQ	258		
DB	555 YVKKKEEYVA--SMGKEMNRLDMLKRTKEEADAAANKKEAQTDLSLKEVEEETIVYLQETVG	613		
QY	259 SADKEIMSLKKKL-----TMLQETLNLPPV---ASETVRLVLESAPAYE	300		
DB	614 EAKAESMMLKENLNDKETRFQNVIHENEDLRAKEDVSLKLIELSKLLEALALAKORPEE	673		
QY	301 VNLKLRPSFRDDIDDLNATFDVDPAPRPSOHHGYEKLECKS-----HSPIDQVP	353		
DB	674 ENGELSE--EKDYDL-----LPKVVEFSSNGCH--RSVEBKAQAVETLDHPPQE--	720		
QY	354 KKKICGPKKESQSLSGGOSGACGEPEELVGAFPIIVRNAILGQKQPKPRRSSSCSKDYV	413		
DB	721 -----QISMGNSNGNGMEKEEVNG-----KPEVETKEKKEKDESODDDKDDSV	763		
QY	414 RTGF 417			
DB	764 EVIF 767			


```
QY      290 RLVLESPAPVENVNKKLRPRFRDDIDLNATFDVDPPARPSSQHGYEKKCLEKSHSPI 349
          :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      431 STTLRFOSLSEKLEKKRKSTERYT-VETAKMLLT-----AKRCVEMKKER 476
          :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY      350 QDVPAKICKGPKRESQLSGSCAGEPDEELVGAFPIFVENAILGOKPRRPSSSC 408
          :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      477 DEMAKE-----KEVEKKLTGGVREEKEKER-----LMETLLGLGEKKRREAIROLIC 523
          :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT   35
A59234    slow myosin heavy chain 3 - quail
C:Species: Coturnix coturnix
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 08-Sep-2000
C:Accession: A59234
R:Nikvolts Jr., W.; Wang, G.F.; Feldman, J.L.; Miller, J.B.; Wade, R.; Nelson, L.; Stock
J. Biol. Chem. 271, 1704-I7056, 1996
A>Title: Isolation and characterization of an avian slow myosin heavy chain gene expressed
A:Reference number: A59234; MUID:96291845
A:Accession: A59234
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1931 <NT>
A:Cross-references: GB:U03862; NID:g1289513; PIDN:AAC59912.1; PID:g1289514
Superfamily: myosin heavy chain; myosin motor domain homology
F:81-761/Domain: myosin motor domain homology <MNO>
```

Query Match	7.0%	Score 166.5;	DB 2;	Length 1931;
Best Local Similarity	21.0%;	Pred. No. 0.43;		
Matches 75; Conservative	74;	Mismatches 131;	Indels 77;	Gaps 11;

```

0Y 1A0EEENWLDREFFLKNELDNVNAQISO - K0KEKDSOYI-----IDTLRDTLEE- 113
Db 1258 IEE5ORTVJDLSTORAKOTIENSELSRQEEKEAFINOLMRGKLTYYTOOLEKRLQEEB 1317
0Y 114 ---RNATVYSILOAQAGKAMELCSLTKKOKYLEDOODETKOQ0EAGRGRSRMKT----- 165
Db 1318 AKAKNALAHALOASACHDCLLEOYEEEBEAKAEI0RALSKNSN5EVA0MRTYETEDAI0R 1377
0Y 166 ---ME0IELL0SOLPEVEEMIRDMGVGSAVE0LAVYCVSLKT-----EYENLK--- 212
Db 1378 TEELEAKKKLQRI0EAE-----AVEANVACSSLEKTKHRLQNETEDIMADV 1427
0Y 213 EARKASGEVADRLKRD0ESSRSKLT0YVSELDQAKLELKSADK0DLOSADKEITMSLKKUL- 271
Db 1428 ERSNMAAALADKK0KNRFDKILSEWK -QKPEESQTELEASQEKARSJSTELFKLNMAE 1484
0Y 272 -----IMLOETLVPYASETYDRLVLESAPYEVNIKLRPRSFRRDIDINATFVDVT 324
Db 1485 ESLEHLEFFKRBNKNLOEISD0LDTOLGASOKSHLE -KVKK0LDAEKLLEQA----- 1537
0Y 325 PRARSSQHGUYEKLCELESHPIDQVYKCI0CKGRKKSQJSLSG0QSAGBPEEL 381
Db 1538 -----ALEEAFASLEHEEGKLTIR0LENNQVKA0EYERKLAKDEEM 1578

```

RESULT 36
A49545
plasminogen-binding protein PAM precursor - Streptococcus pyogenes (fragment)
N:Alternate names: plasminogen-binding M-like protein (Pd 53)
C:Species: Streptococcus pyogenes
C:Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 26-Aug-1999
C:Accession: A49545; S61084; S60829; S70459; S32619
R:Berger, A.; Sjoerding, U.
J. Biol. Chem. 268, 25417-25424, 1993
A:Title: PAM, a novel plasminogen-binding protein from Streptococcus pyogenes
A:Reference number: A49545; MUID:94064605
A:Accession: A49545
A:Molecule type: DNA
A:Residues: 1-388 <BER>
A:Cross-references: EMBL:Z22219; NID:g288978; PIDN:CAA80222.1; PID:g940870
R:Whitmore, A.; Kapur, V.; Sullivan, D.; Musser, J.; Kehoe, M.

submitted to the EMBL Data Library, July 1994

A:Description: Noncongruent relationships between variation in emml gene sequences and

A:Reference number: S61072

A:Accession: S61084

A:Molecule type: DNA

A:Residues: 13-96 <EMHA>

A:Cross-references: EMBL:U11975; NID:9533627; PIDN:AAA99591.1; PID:91235829

A:Experimental source: serotype M53

R:Whitmore, A.M.; Kapur, V.; Sullivan, D.J.; Musser, J.M.; Kehoe, M.A.

Mol. Microbiol. 14, 619-631, 1994

A:Title: Non-congruent relationships between variation in emm gene sequences and the

A:Reference number: S60784; MUID:95198537

A:Accession: S60829

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 17-77 <EMHV>

A:Cross-references: EMBL:U11975

A:Experimental source: serotype M53

R:Carlsson Wistedt, A.; Ringdahl, U.; Mueller-Esterl, W.; Sjoebring, U.

Mol. Microbiol. 18, 569-578, 1995

A:Title: Identification of a plasminogen-binding motif in PAM, a bacterial surface pr

A:Reference number: S70457; MUID:96342385

A:Accession: S70459

A:Molecule type: DNA

A:Residues: 30-162 <CAR>

C:Superfamily: M5 protein

C:Keywords: cell wall

F:1-29/Domain: signal sequence (fragment) #status predicted <SIG>

F:30-388/Product: plasminogen-binding protein PAM #status predicted <MOT>

Query Match	7.0%	Score 165;	DB 2;	Length 388;
Best Local Similarly	20.2%	Pred. NO. 0.066;		
Matches 73; Conservative	62;	Mismatches 120;	Indels 106;	Gaps 11;

[illegible]

RESULT 37
546489
M1 protein precursor - Streptococcus pyogenes
C:Species: Streptococcus pyogenes
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999
C:Accession: S46489; S46490
R:Accession, P.: Schmidt, K.H.; Cooney, J.; Bjorck, L.
Biochem. J. 300, 877-886, 1994
AltTitle: M1 protein and protein H: IgG-Fc- and albumin-binding streptococcal surface p


```

Db 1514 EQLSSSGKTHLEL-KVRKOLEFAEKLELOS-----ALESEAS 1550
OY 349 IODVPRKICKGPKRESQLSIGGOSGAGEPDEL 381
Db 1551 LEHEEGKILNALEFNOIKAKEMERKLAEKDEEM 1583

RESULT 40
A:Accession: A40937
bulbous pemphigoid antigen 1 precursor - human
N:Alternate names: 230k bulbous pemphigoid autoantigen; BP230 antigen
C:Species: Homo sapiens (man)
C:Date: 17-Jul-1992 #sequence, revision 11-Jul-1997 #text change 21-Jul-2000
C:Accession: I56317; A40937; A40882; A61393; A37342; S46669
R:Plgnt G.W.; Stanley, J.R.
J. Invest. Dermatol. 101, 244-246, 1993
A:Title: Cloning of the 5' mRNA for the 230-kD bulbous pemphigoid antigen by rapid amplif
A:Reference number: I56317; M0ID:93346806
A:Accession: I56317
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2649 <RES>
A:Cross-references: GB:LI1690; NID:q402479; PIDN:AAA52288.1; PID:q403124
R:Sawamura, D.; Li, K.; Chu, M.L.; Uitto, J.
J. Biol. Chem. 266, 17784-17790, 1991
A:Title: Human bulbous pemphigoid antigen (BPAG1). Amino acid sequences deduced from cld
A:Reference number: A40937; M0ID:92011453
A:Accession: A40937
A:Molecule type: mRNA
A:Residues: 1-850, 'G', 852-1643, 'T', 1645-2363, 'T', 2365-2494, 'V', 2496-2542, 'K', 2544-2649 <
A:Cross-references: GB:M69225
R:Franka, T.; Parry, D.A.D.; Klaus-Kovtun, V.; Steinert, P.M.; Stanley, J.R.
J. Biol. Chem. 266, 12555-12559, 1991
A:Title: Comparison of molecularly cloned bulbous pemphigoid antigen to desmoplakin I co
A:Reference number: A40882; M0ID:91286285
A:Accession: A40882
A:Molecule type: mRNA
A:Residues: 555-2649 <TRAN>
A:Cross-references: GB:M63618; NID:q179518; PIDN:AAA35606.1; PID:q179519
R:Sawamura, D.; Li, K.H.; Nomura, K.; Sugita, Y.; Christiano, A.M.; Uitto, J.
J. Invest. Dermatol. 96, 908-915, 1991
A:Title: Bulbous pemphigoid antigen: cDNA cloning, cellular expression, and evidence for
A:Reference number: A61393; M0ID:91258872
A:Accession: A61393
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1722-2649 <SAM2>
A:Note: the sequence from Fig. 3 is inconsistent with the nucleotide sequence from Fig.
R:Owaribe, K.; Kartenbeck, J.; Stump, S.; Magin, T.M.; Krieg, T.; Diaz, L.A.; Franke, W
Differentiation 45, 207-220, 1990
A:Title: The hemidesmosomal plaque. Characterization of a major constituent protein as a
A:Reference number: A37342; M0ID:91216368
A:Accession: A37342
A:Molecule type: mRNA
A:Residues: 1595-1942, 'R', 1944-2202 <OMA>
A:Cross-references: GB:X58677; NID:q36094; PIDN:CAA41528.1; PID:q36095
R:Hopkinson, S.B.; Jones, J.C.R.
Biochem. J. 300, 851-857, 1994
A:Title: Identification of a second protein product of the gene encoding a human epiderm
A:Reference number: S46669; M0ID:94280413
A:Accession: S46669
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1897-2081 <HOP>
A:Cross-references: EMBL:U04850; NID:q451556; PID:q451557
C:Genetics:
A:Gene: GDB:BPAG1
A:Cross-references: GDB:125207; OMIM:113810
A:Map position: 6p12-6p11
A:Introns: 2016/3
A:Note: the list of introns is not complete
C:Keywords: basement membrane; extracellular matrix; glycoprotein; membrane protein; pol

```

F:1-43/Domain: signal sequence #status predicted <SIG>

Query Match 7.0% Score 166; DB 2; Length 2649;
 Best Local Similarity 20.0%; Pred. No. 0.66;
 Matches 97; Conservative 88; Mismatches 181; Indels 120; Gaps 19;

```

OY 33 CLIQSFETAPSRPCPCRIQGVKRTIINKLFFDLAEEENVLDREFLNKELNDVNR----- 88
Db 1330 CRENALPVCPTIQTQSCRAVTG-----LQGEHDKQAEELKQGVDELTPANKR 1377
OY 89 -----QLSQDKERKDSQVIIDTLRLTLERNATVVSLOALGAEMLCSTLKRO 138
Db 1378 AEQDMRELFLYELNALOLEKTSSEKARLLKDKLDENNTFLRLCLLEKRDQAEKYSQO 1437
OY 139 MKYLEQOQDET-----KQAGEACRLSKKMTMQIEL-----LLQSLPVEEMIRD 186
Db 1438 LRELGRQNLQTTGKAEAMQEAASDLK-KIKRNYOLEESLNHEKGKIQR---EVDRIITRA 1493
OY 187 MGVGQSAVEOL--AVYCVSLKREYENLKEARKKASGEVADLRKDLFSSRSKLTQVYSELD 244
Db 1494 HAVAEKRIQHLNSQIHSPREKELERLQICRKSDDLKEQFER---SHEOLLONITAEKE 1550
OY 245 QAKLELKSAGKDQSADKEIMSLKKLTMLQETLNLPPVASETV---DLVLESBPAYE 300
Db 1551 NND-KIQRLNEBELKESMECAEMLKQVEELTRQNNETKLMQRIQAESENIYVEKQTIQ 1609
OY 301 --VNLKLRPSFDDIDLNATFPVDTPPARPSSQHGTYEKLCLKLEKSHSPIDVPPKICK 358
Db 1610 RCEALKTQADGFPDOL-----RSTNEH-----LHKQTKEDQDFORKI-- 1646
OY 359 GPKRESQLSIGGOSGAGE---PDELVGAPP-----IFVRNALIGOKQPKRPRSSSC 408
Db 1647 -----KCLEEDLAKSONLVSEFKQCDQONITIIQNT---KKEVRNLMDELNA 1690
OY 409 SKDVVRITGFDGLGRTKFTIQTPTVTMIRPLPVKPKTKQKQKRVKVTVP-----LFG 460
Db 1691 SKEEKRRGEGQKVOLOQAOVQELNNRL-----KKVQDELHLKTEEDQMTIRKKNVLFQ 1741
OY 461 AKLDTF 466
Db 1742 EESGKF 1747

```

Search completed: September 4, 2002, 16:12:45
 Job time: 3229 sec

.....
